The Practice of Informatics

Electronic Support for Public Health: Validated Case Finding and Reporting for Notifiable Diseases Using Electronic Medical Data

ROSS LAZARUS, MBBS, MPH, MICHAEL KLOMPAS, MD, MPH, FRANCIS X. CAMPION, MD,
SCOTT J. N. MCNABB, PHD, MS, XUANLIN HOU, MSC, JAMES DANIEL, MPH, GILLIAN HANEY, MPH,
ALFRED DEMARIA, MD, LESLIE LENERT, MD, MS, RICHARD PLATT, MD, MSC

A b s t r a c t   Health care providers are legally obliged to report cases of specified diseases to public health authorities, but existing manual, provider-initiated reporting systems generally result in incomplete, error-prone, and tardy information flow. Automated laboratory-based reports are more likely accurate and timely, but lack clinical information and treatment details. Here, we describe the Electronic Support for Public Health (ESP) application, a robust, automated, secure, portable public health detection and messaging system for cases of notifiable diseases. The ESP application applies disease specific logic to any complete source of electronic medical data in a fully automated process, and supports an optional case management workflow system for case notification control. All relevant clinical, laboratory and demographic details are securely transferred to the local health authority as an HL7 message. The ESP application has operated continuously in production mode since January 2007, applying rigorously validated case identification logic to ambulatory EMR data from more than 600,000 patients. Source code for this highly interoperable application is freely available under an approved open-source license at http://esphealth.org.

Introduction and Background

Mandatory reporting of infectious diseases forms a cornerstone of cost-efficient, preventive public health programs. Timely, complete and accurate case reports facilitate contact tracing, appropriate treatment, and follow-up, decreasing the risk of an infection being spread.[1,2] In the United States, voluntary systematic disease reporting by physicians dates from 1874.[3] Reporting of patients found to have specific diseases such as common sexually transmitted infections, as well as less common infections like tuberculosis, is now a legal requirement for health care practitioners in all U.S. states. Unfortunately, despite substantial progress, including the introduction of electronic reporting forms,[4,5] many of these important surveillance mechanisms still depend on practitioner initiated, manual data entry and submission.

Busy clinicians find it frustrating and burdensome to manually transcribe clinical and demographic patient details between independent data systems. Not surprisingly, most evidence suggests that practitioner initiated, manual reporting systems provide delayed[6–8] and inaccurate data, with many omissions and errors.[9] Reporting systems based on laboratory test results are increasingly able to supplement practitioner initiated reporting, potentially detecting many more cases and reporting with less delay,[10] but are not always able to provide complete reporting[7,11] and cannot identify conditions defined by clinical criteria, such as acute pelvic inflammatory disease (PID). More importantly, laboratory reporting systems lack crucial clinical details[6,11,12] for public health practitioners involved in managing reported cases, such as vital signs, drug, dose, and route of antibiotic or other treatment, and in the case of a female patient, pregnancy status.

We based the work described here on the premise that if suitable data are available in electronic form, an appropriate software application could support automated detection of notifiable diseases, facilitating the timely reporting of cases, including all relevant clinical information, without requiring practitioner initiation or error-prone manual data transcrip-

Affiliations of the authors: Department of Ambulatory Care and Prevention, Harvard Medical School and Harvard Pilgrim Health Care (RL, MK, XH, RP), Boston MA; Channing Laboratory, Brigham and Women's Hospital (RL, MK, XH, RP), Boston MA; Atrius Health (FXC), Boston, MA; National Center for Public Health Informatics, Centers for Disease Control and Prevention (SJNMN, LL), Atlanta, GA; Massachusetts Department of Public Health (JD, GH, ADM), Boston, MA.

Correspondence: Dr Ross Lazarus, Channing Laboratory, 181 Longwood Ave., Boston, MA 02115; e-mail: <ross.lazarus@gmail.com>.

tion. Such a system could be adapted to work with data integrated from diverse sources of electronic health data, such as regional health information exchanges, the electronic medical records of large medical practices, and, for more limited purposes, laboratory reporting systems, or pharmacy benefits managers.

The Electronic Support for Public Health (ESP) application[13,14] is an automated, platform independent disease detection and reporting system, developed in an ongoing collaboration between the Centers for Disease Control and Prevention (CDC), the CDC funded Center of Excellence in Public Health Informatics based at Harvard, Harvard Vanguard Medical Associates (HVMA) and Atrius, and the Massachusetts Department of Public Health (MDPH). The initial installation of ESP uses a fully automated data flow from an integrated commercial EMR system used by more than 700 physicians spread over 30 practice sites, serving more than 600,000 patients, providing near real-time notifiable disease case detection, and secure, standards-based, automated, electronic communication to the relevant public health authority. We have previously outlined some aspects of the planning and implementation of the ESP project.[14.] In this report, we describe key informatics issues and features, and summarize nearly two years of continuous, live ESP operation, including potentially useful lessons learned from this effective, interoperable, and extensible public health informatics application.

Design Objectives

Motivation
Completing statutory disease reporting forms by hand requires transposing patient and clinical details from one system into another, in an inefficient and error prone process that diverts valuable time from a busy clinical schedule. Portable hand held and web-based systems can replace paper forms with a modern electronic equivalent,[4,9] but that investment remains dependent on practitioner initiated, manual data entry. The ESP system was designed to automate the identification of valid cases from an appropriately comprehensive electronic data stream, and to provide secure, automated, detailed reporting for patients who satisfied specific diagnostic and other criteria.

Proposed Evaluation Criteria
The ideal automated notifiable disease detection and reporting system would be perfectly secure, have perfect validity and precision, consume minimal resources, be readily portable to other data streams and other messaging specifications, support multiple useful functions, be freely distributable, and run on commodity technology to minimize marginal costs. In this section, each of these ideal, but unattainable in practice criteria is described in more detail, and the extent to which the operational ESP system meets these criteria is reviewed in the Discussion.

Security
Patient identifiers are required for case reports, so security was the most fundamental design imperative. New applications like ESP carry potential risk exposure for host organizations, in addition to potential benefits. We addressed this challenge by adopting an open-source, distributed model,[15–17] because it overcomes many potential concerns about securing identifiable EMR data against inadvertent disclo-

sure. In a distributed approach, the application software goes to the data rather than moving the data to a central application site for analysis.[18] An independent, ESP server can be installed wherever the host EMR servers reside. This design serves to minimize any increased risk of inadvertent disclosure of identifiable patient data, allowing the host organization to retain complete control over all access to the system and data, and all ESP source code is available for scrutiny.

External Validation—Sensitivity and Specificity
False positive case reports waste valuable effort, while failure to report a real case (false negative) will decrease the effectiveness of the preventive public health program. A major design goal for ESP was to achieve reporting sensitivity and specificity that were as close to perfect as practicable. Substantial resources were devoted to the collaborations, and to the painstaking manual validation processes needed to quantify both false positive and false negative reporting error rates. Measures from this external validation process were regarded as the key figures of merit for refining the case detection algorithms for each reportable condition. While electronic laboratory reporting systems might be assumed to be the "gold standard" for timely disease surveillance and reporting,[10] previous reports[11] have indicated that they are not always perfect, and they lack crucial information required for detecting cases that include clinical criteria, such as PID, or where a chronic infection must be reliably distinguished from an acute one.

Completeness of Reports
Complete, detailed case data is needed for public health intervention, but is generally not available from manually prepared standard notification forms. An automated system based on a comprehensive source of electronic data can provide whatever details were available in the incoming data stream, potentially making the job of managing reported cases easier by avoiding the need to obtain or transcribe additional data. Ideally, "completeness" could be quantified as the proportion of required data items available in the data feed, and present in case reports, if agreement could be reached on which data items are "required" for each type of case, and we are currently exploring this as a new metric for future research.

Resources
The capital costs of software and hardware are minimized by using open-source software, and efficient software engineering techniques, so that inexpensive, readily available hardware will be adequate, even for very large EMR volumes. Installation will always require some local tailoring, and all local code changes are isolated to two small modules and have no impact on the rest of the system. This decoupled, modular design was chosen to isolate all effort for portability. Secure remote administration helps keep running costs low, and appropriately designed administrative interfaces allow local staff to perform most routine maintenance, further minimizing recurrent costs. The development, testing and validation of case detection criteria and logic, and the successful negotiation and testing of standards for report message formatting, security, and transmission, all require effective partnerships, such as the longstanding and highly successful collaboration between the Harvard

Medical School Department of Ambulatory Care and Prevention (DACP), HVMA and MDPH.

Utility

Maximizing the range of useful functions supported by the application increases its utility, making the business case for adoption more attractive. The ESP application has been designed and built as a generic framework, with multiple potential uses in mind. Given the established infrastructure for data flow and notifiable disease case reporting, additional functions, such as vaccine adverse event detection and reporting currently under development, or other practice quality assurance activities, can now also be relatively easily sustained, using the ESP database tables.

Interoperability and Portability

Substantial effort was required to develop, refine and validate case identification rules for a reliable automated system. The total effort to build and validate one case identification system that is reliably transportable to any other data source, is probably far less than the total effort required to build and validate a new system for every new data stream. For a case detection system to be portable, it must be adaptable, so interoperability with other electronic data sources, and public health systems, was a major design goal. While data format and other relatively low-level standards are often a focus for discussions of interoperability, valid, code-based case detection logic is only reliably inter-operable when there is uniform internal consistency in codes across all instances. For a system like ESP, each independent instance must be robust when faced with locally tailored and constantly changing coding systems for laboratory test orders and results, diagnostic codes, diagnoses and medications.

System Description

Architecture

The ESP system has a very loosely coupled, modular design, conceptualized as a set of independent software components that communicate through simple, explicit interfaces. All external communications are confined to two relatively low-complexity, pluggable interface modules (see Figure 1),
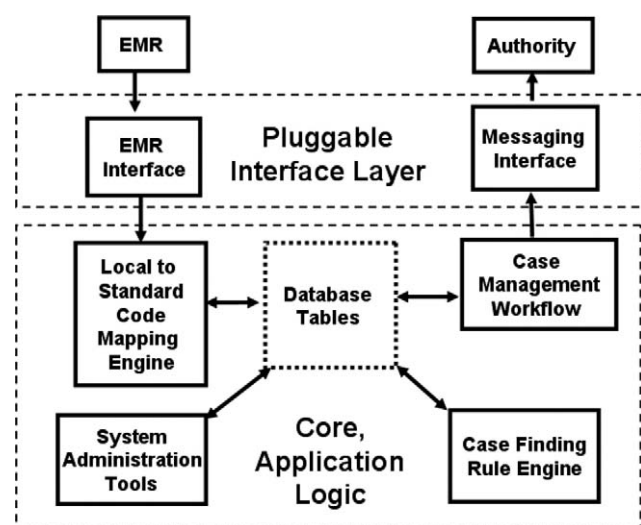
Figure 1. ESP System Architecture.

tailored to fit variable incoming local data streams and outgoing public health authority messaging interfaces respectively. The core modules are all designed around a relational database schema, where patient data, notifiable cases, and workflow states are managed and stored, in addition to case definition criteria and other internal application tables. Core modules include an automated rule engine that reads patient data and creates new cases or updates existing cases, and an optional, interactive case management workflow module. Modules were designed to operate independently, and all non-interactive component collaborations are sequenced using scripts running automatically each day. Interactive, web-based administrative tools were built to manage application security, manage system and instance configuration tables and to extract data for case validation, report auditing and logging.

The ESP system code that may require alterations to conform to specific local standards at each independent installation is restricted to the two external, pluggable interface "border" layers illustrated in Figure 1. Replacing these with any functionally equivalent component, has no effect on the other, core modules.

Physical and Logical Security

To attain the required levels of data protection and system security, all ESP software is made available as source code, designed to run with minimal support on a dedicated server, located in the host health care provider's computing center, enabling the ESP system to be secured by physical and other measures already in place to protect identifiable patient data. In addition, this design feature helps to isolate any computational and storage load from interfering with the host production systems, making it more acceptable to host organizations.

The EMR Interface and Code Mapping

External data are loaded into ESP for processing (Figure 1, upper left). *Epic Care* from Epic Systems, Inc, (Verona, WI) was the source data stream for ESP development and initial deployment, and serves as a model and proof-of-concept for other commercial EMR systems. A locally tailored extract, transform and load (ETL) procedure provides a periodic incremental delimited text file extract from the host EMR. As these text files become available, ESP completes the transform and load steps. Alternatively, an HL7 interface using the open-source Mirth project gateway (http://mirthproject.org) has also been implemented.

The ESP application was designed to operate at multiple sites with a single validated set of case criteria operating over a uniform set of specific diagnostic and other codes for each condition. No matter how ESP tables are loaded, some local codes will probably need to be converted into the specific LOINC, SNOMED, and ICD codes expected by validated ESP notifiable disease case detection logic. The ESP application provides a module to manage and deploy this code mapping for the local incoming data, initially configured at installation. Note that in practice, this is not as daunting a task as might be thought, because mappings are only needed for the relatively few codes specific to notifiable conditions.

Case Identification and Reporting Logic

In very general terms, logic for each condition is internally represented as rules and sets of definitive codes and other characteristics. A publication by Klompas et al.[19] describes the development of the acute hepatitis B detection algorithm in some detail. Current algorithm specifications, and the source code implementing them, are readily available from the project web site (http://esphealth.org). Briefly, the specific criteria required to define or exclude a case are stored in database tables. A separate list of codes and other characteristics to be included in the notification message is also maintained in database tables. These permit reporting of the specific data elements required for each condition, while protecting other confidential information. Our MDPH collaborators were adamant about specifying which data elements are reported for each category of notifiable disease, because they are only authorized access to identifiable data directly relevant to specific public health purposes. Code sets for each condition are read and used when the case identification logic is run.

Natural Language Processing

In the real world, electronic health record systems contain diverse, and constantly changing code "ecologies", as suppliers of laboratory services, and medical practices deploy new equipment and tests, leading to potentially important changes to their coding systems over time. Simple natural language processing using text string regular expression matching is applied to all incoming codes, to identify changes of potential importance in the data feed system. Any text that might indicate a new test code for an organism of interest (e.g., "gonor*" to match any text related to *Neisseria gonorrhoeae*) discovered in each day's new test results is automatically emailed to the administrative staff for appropriate action. As new relevant codes are discovered, they are manually added to the translation tables and all existing instances of the new codes are automatically transformed.

Message Generator and Message Transport for Approved Cases

Case reporting message format specifications were based on the existing MDPH Electronic Laboratory Reporting (ELR) HL7 specification, in turn based on the CDC ELR specification. The message generator and transport interface specific to this specification is distributed with ESP source, where it serves as a flexible prototype, but it is easily replaced with any generator compatible with the underlying database structures.

Status Report

Current Operational Status

The ESP system began daily operation in January 2007, and all data from July 1, 2006 were "backfilled" to facilitate external validation described below. Since then, ESP has reported more than 1490 cases of chlamydia, 196 cases of gonorrhea, 31 cases of PID, six cases of acute hepatitis A, 10 cases of acute hepatitis B, six cases of acute hepatitis C and 13 cases of active tuberculosis.[20] The ESP application currently reports cases for more than 700 physicians spread out over more than 30 practice sites within the Harvard Vanguard-Atrius Health (http://atriushealth.org) system. Priorities for implementation of specific diseases have been driven by the needs of our clinical and MDPH partners, with particular emphasis on public health importance, and perceived under-reporting in other existing reporting systems.

External Validity—Sensitivity and Specificity

To check for missed true cases, and for false positive cases, all historical cases reported manually by an independent dedicated team, and those collected by the State from all sources for Atrius patients, were manually compared with cases reported for the same period by ESP logic, as described in more detail elsewhere.[21] In summary, 758 cases of chlamydia, 95 cases of gonorrhea, 20 cases of pelvic inflammatory disease, and four cases of acute hepatitis A were detected by ESP in the 12 months to July 2007. Manual review of all case charts and comparison with all conventional reports received by MDPH revealed that ESP reported more cases (758 cases versus 545 for chlamydia, 95 cases versus 62 for gonorrhea, 20 versus zero for PID, four versus one for acute hepatitis A). Six traditionally reported cases of chlamydia were not detected by ESP, of which five were false positives. The single true case missed had been assigned an incorrect laboratory test code in the host EMR system. No cases of gonorrhea, pelvic inflammatory disease, or acute hepatitis A detected by passive surveillance were missed by ESP. So, sensitivity specificity and positive predictive value were all close to ideal, using the current health department data as the "gold" standard, and in fact, many more real cases were identified.

Case Report Completeness and Transcription Errors

In external validation,[21] conventional reports noted pregnancy status for only 5% of female cases and treatment status for 88% of all cases compared to 100% for both in ESP reports. Patient name spelling errors were detected in 5% of conventional manual reports when compared to ESP reports derived directly from electronic administrative and clinical data, which we assume contain the correct patient details.

Natural Language Processing

The Harvard Vanguard-Atrius Health deployment of ESP receives coded and free text laboratory results from six distinct providers (five group practices and a major private laboratory), each of which is at liberty to make changes at any time to their reporting practices, texts and codes, without any formal notification to the ESP staff. Most of these changes make no difference to notifiable disease case detection, but ESP runs an automated natural language surveillance process for code changes, that generates alerts when new codes of potential interest are detected. For example, the alerting system responded appropriately and immediately, during two unscheduled tests during 2007, when new clinical settings and group practices were added to the ESP data feed without prior warning from the EMR operations staff. The ESP application was quickly reconfigured to deal with all new, relevant codes using the management interface, and all relevant potentially false negative cases were correctly detected and reported once the relevant codes had been added to the database tables.

Throughput and Stability

The host EMR system averages 12,000 ambulatory care encounters each day from approximately 600,000 patients. The various ESP database tables have accumulated a total of

more than 60 million rows, including 18 months of prospectively collected records, and 6 months of "backfilled" historical records. A very modest server configuration (Sun X2100 with 2GB RAM and a dual core Opteron processor) is adequate for this load and, as is typical of Linux systems in our experience, has operated continuously, without requiring "rebooting", for more than 18 months since deployment.

Discussion

Ensuring appropriate treatment and preventing spread of an infection is a highly effective public health strategy, but it depends on cases being detected and brought to the attention of appropriate public health agencies in a timely manner. Manual surveillance systems may be the only alternative for manual medical record systems, and electronic laboratory based systems may substantially improve case finding compared with manual systems.[10] Our experience[20] suggests that comprehensive electronic clinical data can support highly sensitive and specific surveillance, yielding substantially improved compliance with statutory obligations over manual methods. An evaluation in terms of the stated ideal design objectives follows below.

Security

The installation of ESP is effectively as secure as the host EMR system itself. The distributed open-source model minimizes any marginal increase in risk of inadvertent disclosure of identifiable patient information.

Validity and Performance Characteristics

The ESP system was externally validated and found to perform extremely well.[20] It is robust and reliable in production, able to detect and withstand changes in EMR codes of interest when unanticipated systematic changes occur.

Resources

The ESP application was a CDC-funded academic research and demonstration software development project with a very modest budget compared to commercial undertakings of comparable scale and complexity. Code and case detection rule development continues with CDC funding, and these deliverables are freely distributed as part of our funded research. In our experience, successful implementation of a new ESP instance required four to six weeks effort from dedicated local administrative and EMR technical staff for installation, and although largely automated, a few hours a week of dedicated, ongoing administrative and remote technical support. This effort is required whenever a new interface module is created to ensure that the electronic data feed is complete, reliable and accurate; that all appropriate codes are correctly mapped; that the system has an acceptable sensitivity and specificity which can only be quantified by checking with existing manual reporting systems; and that the system meets the operational requirements of the health department concerned—in the case of MDPH, a four week period of reliable operation in test mode was mandated before the system was certified for production use. If ESP were "bundled" with an EMR or electronic laboratory reporting system, nearly all of this effort would be avoided for multiple installations, after validation of an exemplar instance.

Utility

Tables and relationships in the ESP data model are very general, allowing new software applications to add value to the EMR data held in ESP, at relatively low marginal cost. For example, the ESP team is adding an automated secure vaccine adverse event reporting module to the system described here, in collaboration with the CDC and with support from the Agency for Healthcare Research and Quality (AHRQ). Development of new modules to support quality assurance activities, such as notifying designated practitioners of patients who have not received appropriate follow-up diagnostic studies or treatment within a specified period, or to perform post-marketing surveillance for adverse events from medication, are all now feasible, and the investment in their development carries relatively low business risk, given the existing stable ESP platform, infrastructure and data flows.

Interoperability and Portability

The ESP application was designed to make reliable interoperability and portability as straightforward as possible. The ESP application is currently being installed at the Northern Berkshire e-health Collaborative site (http://www.maehc. org/NorthAdams.html), where an HL7 gateway will be used for incoming EMR data from an eClinicalWorks (http://www.eclinicalworks.com) EMR system, and the existing HL7 messaging module will be used to send notifications to the MDPH once the code mapping tables and validation processes are completed.

Lessons Learned

Our experience in deploying ESP suggests some useful lessons about a generic, distributed, notifiable disease case identification and reporting framework. For a distributed system, the required subsystems are readily enumerated—a flexible incoming gateway for HL7 or ETL data; a reliable mapping for heterogeneous codes to uniform, standard nomenclatures used in the case detection rules; a portable representation for a set of rules; locally tailored message formatting and secure messaging subsystems; and administrative applications to support code mapping maintenance, case management, record keeping, and ongoing validation. Each of these is briefly discussed in turn below.

Distributed Model

In our experience, there is a greater load on the informatics team running a distributed application compared to a team collecting and analyzing data centrally. However there are many benefits from the distributed model, including increased willingness of data custodians to collaborate, because they retain control over identifiable patient data and associated security risks. An additional and particularly valuable resource available to a distributed system is the intimate local knowledge and insight from local EMR staff, and from dedicated case managers and administrators, at installation and in ongoing mapping local EMR codes to standard codes. Secure remote system administration, source code revision management and distribution to remote production systems could potentially be performed by a single, specialized team managed by a public health authority, or commercial application vendor.

Incoming Data

The initial ESP deployment uses a conventional extract, transform, load (ETL) process, implemented in collaboration with the host EMR programming staff. This model was adopted because it has been proven in a multi-site, distrib-

uted national bioterrorism surveillance system that we deployed and operated successfully for nearly seven years.[16,18] The local EMR staff manages a periodic process that extracts all transactions in the previous period from the host EMR into delimited text files. Resulting periodic text files are made available for the application to process. Currently, the period used is 24 hours, chosen to fit the operational requirements of the host EMR system, but more frequent timing could be used if required. An incoming HL7 gateway, using the open-source Mirth HL7 server (http://mirthproject.org) was added for the Northern Berkshire e-health Collaborative ESP deployment, and this use of an HL7 listener as the source of incoming EMR data could facilitate near real-time reporting for time-critical situational awareness applications, as demonstrated in RODS[21] and in the CDC Biosense[22] network. The benefits and costs of an HL7 gateway implementation compared with the ETL approach will be the subject of future research as we gain additional experience.

Practicalities in Mapping Local EMR to Uniform Codes

A key practical barrier to creating readily transportable case identification applications, is that available EMR software implementations have idiosyncratic, heterogeneous and constantly changing coding systems. In contrast, a portable, validated case detection system relies on absolutely specific codes. The issues raised by diversities in system vocabulary are more subtle, but equally applicable between sites using the same software system, since codes are typically modified locally to suit the practice staff and their testing laboratories.

In the initial ESP deployment site, laboratory tests are identified using CPT codes, and laboratory test results are generally only available as text, such as "Positive" or "Not detected". The HL7 specification required by the MDPH requires specific LOINC codes and SNOMED codes for each laboratory test result reported. We chose to map between the codes and text in the incoming EMR data before the data is stored in internal database tables, and built a substantial infrastructure to perform this mapping in an automated manner using mapping tables also stored in the database. These mapping tables have a separate administrative interface accessible to authorized users through the case management website, to help ensure that a single uniform set of case detection rules operates correctly at each independent ESP installation. With an appropriate security infrastructure in place, this architecture allows rules to be updated remotely, and this could be used to support coordinated response to evolving public health emergencies.

Representation and Implementation of Validated Rules

Uniform internal representation across multiple ESP instances seems a potentially efficient approach to ensuring that a single agreed set of case definition rules for each notifiable condition can be represented in a transportable way to multiple EMR systems. For performance and other practical reasons, the representation of codes, rules and logic in ESP used table-driven code sets, with additional logic for complex cases expressed in a high-level language.

Installation and Dedicated Ongoing Costs

Although there will never be software licensing costs for ESP, installation effort will vary with local requirements, being lowest for completely compatible data streams and messaging formats as the current sites. Each independent installation will involve some dedicated effort for validating the incoming data gateway, code mapping, messaging and case validation processes, before acceptance testing and production operation can commence. The complexities of EMR based case detection mean that maintaining a system like ESP will require ongoing dedicated effort. Local staff with ordinary web-browser access can configure and manage the local code mappings to ensure that all incoming codes needed for the case detection logic are reliably translated. While the design of ESP minimizes this effort, there are inevitable maintenance costs for ESP, so it must offer direct benefits to health-care organizations installing it or other public health reporting systems. The ESP:VAERS application is a parallel project to report adverse events after vaccination from the same data stream.

Unintended Consequences

It is possible that increased sensitivity to true cases from a comprehensive automated system may lead to an increased volume of cases. If the additional volume from an automated system is substantial compared to existing manual reporting, the additional workload on the public health authority may require additional resources and changes to existing systems. Informal feedback from MDPH staff has been uniformly positive, with the additional case load being at least somewhat offset by highly reliable and complete patient and clinical data contained in the automated reports.

Conclusions

Automated notifiable disease case finding and secure reporting systems are practicable, adding value to existing medical record data streams, and leading to improvements in the completeness, timeliness and accuracy of reporting compared with existing manual systems, providing more complete clinical data than electronic laboratory reporting systems. Designing for portability requires easily reconfigurable external input and output interfaces, and on infrastructure to ensure that each independent system instance correctly translates codes needed for validated logic from incoming EMR data. Dedicated effort and vigilance supported by simple natural language processing is essential to detect changes in the ecology of test and other important codes, in order to ensure reliable and valid ongoing operation. A source code repository and other resources are available at http://es-phealth.org.

References ∎

1. Chorba TL, Berkelman RL, Safford SK, Gibbs NP, Hull HF. Mandatory reporting of infectious diseases by clinicians. J Am Med Assoc 1989;262(21):3018–26.
2. Roush S, Birkhead G, Koo D, Cobb A, Fleming D. Mandatory reporting of diseases and conditions by health care professionals and laboratories. J Am Med Assoc 1999 Jul 14;282(2):164–70.
3. Centers for Disease Control and Prevention. Mandatory reporting of infectious diseases by clinicians. Morbidity and Mortality Weekly Reports (MMWR). 1990;39(RR–9):16–7.
4. Lober WB, Bliss D, Dockrey MR, Davidson AJ, Karras BT. Communicable disease case entry using PDAs and public wireless networks. AMIA Annu Symp Proc. 2003:916.

5. Centers for Disease Control and Prevention. Progress in improving state and local disease surveillance—United States, 2000-2005. MMWR Morb Mortal Wkly Rep. 2005 Aug 26;54(33): 822–5.

6. Birkhead G, Chorba TL, Root S, Klaucke DN, Gibbs NJ. Timeliness of National Reporting of Communicable Diseases: The Experience of the National Telecommunications System for Surveillance. Am J Public Health 1991;81(10):1313–5.

7. Effler P, Ching-Lee M, Bogard A, Ieong MC, Nekomoto T, Jernigan D. Statewide system of electronic notifiable disease reporting from clinical laboratories: comparing automated reporting with conventional methods. JAMA 1999 Nov 17; 282(19):1845–50.

8. Jajosky RA, Groseclose SL. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. BMC Public Health 2004 Jul 26;4:29.

9. Ward M, Brandsema P, van Straten E, Bosman A. Electronic reporting improves timeliness and completeness of infectious disease notification, The Netherlands, 2003. Euro Surveill. 2005 Jan;10(1):27–30.

10. Overhage JM, Grannis S, McDonald CJ. A comparison of the completeness and timeliness of automated electronic laboratory reporting and spontaneous reporting of notifiable conditions. Am J Public Health 2008;98(2):344–50.

11. Panackal AA, M'Ikanatha N M, Tsui FC, et al. Automatic electronic laboratory-based reporting of notifiable infectious diseases at a large health system. Emerg Infect Dis. 2002 Jul;8(7):685–91.

12. Backer HD, Bissell SR, Vugia DJ. Disease reporting from an automated laboratory-based reporting system to a state health department via local county health departments. Public Health Rep. 2001 May-Jun;116(3):257–65.

13. Klompas M, Lazarus R, Daniel J, et al. *E*lectronic Medical Record *S*upport for *P*ublic Health (ESP): Automated Detection and Reporting of Statutory Notifiable Diseases to Public Health Authorities. Adv Dis Surveill 2007;3(3):1–5.

14. Klompas M, Lazarus R, Haney G, et al. The Electronic Support for Public Health (ESP) Project: Automated Detection and Electronic Reporting of Notifiable Diseases. Adv Dis Surveill 2007;4:52.

15. Yih WK, Caldwell B, Harmon R, et al. National Bioterrorism Syndromic Surveillance Demonstration Program. Morbidity Mortality Weekly Report Supplement. 2004;53 Suppl.:43–9.

16. Lazarus R, Kleinman K, Dashevsky I, et al. Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. Emerg Infect Dis 2002;8:753–60.

17. Lazarus R, Kleinman KP, Dashevsky I, Alfred DeMaria, Platt R. Using automated medical records for rapid identification of illness syndromes (syndromic surveillance): the example of lower respiratory infection. BMC Public Health 2001;1:1–9.

18. Lazarus R, Yih WK, Platt R. Distributed data processing for public health surveillance. BMC Public Health 2006;19(6):235.

19. Klompas M, Haney G, Church D, Lazarus R, Hou X, Platt R. Automated identification of acute hepatitis B using electronic medical record data to facilitate public health surveillance. PLoS ONE. 2008;3:e2626.

20. Klompas M, Lazarus R, Platt R, et al. Automated detection and reporting of notifiable diseases using electronic medical records versus passive surveillance--Massachusetts, June 2006–July 2007. MMWR Morb Mortal Wkly Rep. 2008;57:373–6.

21. Wagner M, Espino J, Tsui F, et al. Free Full Text Syndrome and outbreak detection using chief-complaint data--experience of the Real-Time Outbreak and Disease Surveillance project. MMWR Morb Mortal Wkly Rep. 2004;24(53):S28–31.

22. Bradley C, Rolka H, Walker D, Loonsk J. BioSense: implementation of a National Early Event Detection and Situational Awareness System. MMWR Morb Mortal Wkly Rep. 2005; 26(54):S11–9.